

DUT Informatique

Année universitaire 2021 / 2022

M4202C – Statistique inférentielle

Responsable : A. Ridard
Autre intervenant : L. Naert



Avant-propos

Ce document est spécifiquement rédigé pour des séances de Cours/TD.

Il présente les éléments de cours habituels (définitions et propriétés) enrichis de remarques, indiquées par , donnant un certain éclairage pour mieux les comprendre, les retenir et les utiliser.

Ce cours est aussi ponctué d'exercices, indiqués par , qui seront traités en classe ou à la maison pour bien assimiler les différentes notions présentées. C'est effectivement en étant acteur dans ses apprentissages que l'on profite au mieux des enseignements!

Ce document sera complété par des feuilles de TD ou TP pour s'entraîner d'avantage.

Bonne lecture, et bon travail...

Table des matières

A Introduction à la Statistique inférentielle	7
I. Statistique descriptive / Statistique inférentielle	7
II. Modes d'échantillonnage	7
1. Sondage aléatoire simple	7
2. Sondage en strates	9
III. Paramètres étudiés	9
1. Moyenne et variance d'une v.a.	9
2. Proportion (hors programme)	10
B Estimation	11
I. Estimation ponctuelle	11
1. Estimateur et Loi Forte des Grands Nombres (LFGN)	11
2. Qualités d'un estimateur	12
II. Estimation par intervalle de confiance	14
1. Principe	14
2. Moyenne	14
3. Variance	16
4. Proportion (hors programme)	17
C Tests d'hypothèses	19
I. Introduction	19
1. Les faiseurs de pluie	19
2. Principe	21
II. Tests de conformité	21
1. Moyenne	21
2. Variance	22
3. Proportion (hors programme)	23
III. Tests de comparaison de deux échantillons indépendants (projet)	23
1. Moyennes	23
2. Variances	24
3. Proportions (hors programme)	26
Annexes	27

Chapitre A

Introduction à la Statistique inférentielle

Sommaire

I. Statistique descriptive / Statistique inférentielle	7
II. Modes d'échantillonnage	7
1. Sondage aléatoire simple	7
2. Sondage en strates	9
III. Paramètres étudiés	9
1. Moyenne et variance d'une v.a.	9
2. Proportion (hors programme)	10

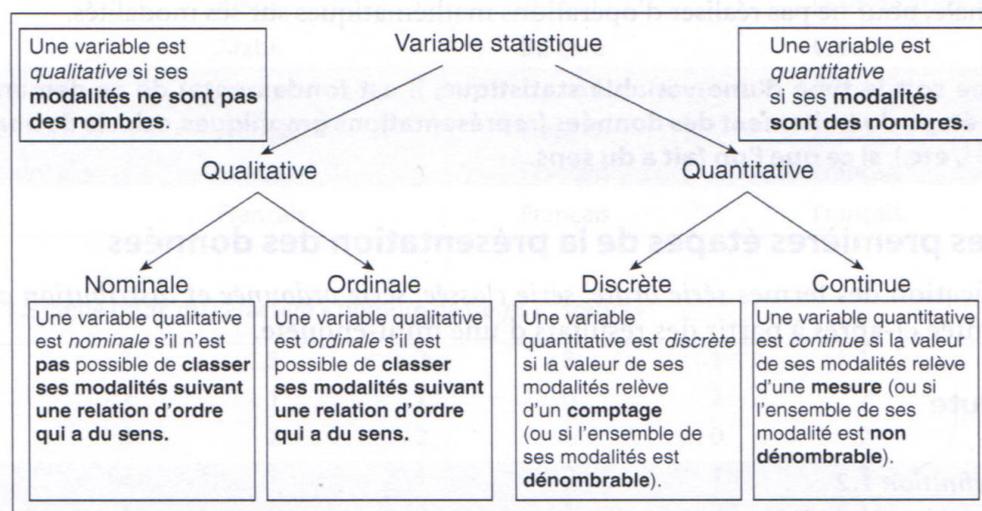
I. Statistique descriptive / Statistique inférentielle

Si la Statistique descriptive consiste en l'étude d'une population toute entière d'individus selon un ou plusieurs caractères, la Statistique inférentielle permet de déduire des informations sur une population de taille N à partir d'un échantillon de taille n . Avant de préciser ce que l'on souhaite inférer, nous allons présenter différentes manières de prélever un échantillon.

II. Modes d'échantillonnage

1. Sondage aléatoire simple

Il est important de rappeler que chaque individu d'une population est caractérisé par un ou plusieurs caractères appelés aussi variables. On distingue deux types et quatre sous-types de variables.



Définition (variable aléatoire).

Une variable (statistique) est une application (au sens mathématique du terme) qui, à chaque individu, associe une valeur (numérique ou non).

Autrement dit, si l'on note ω un individu de la population et X la variable étudiée, alors $X(\omega) = x$ signifie que le caractère X a pour valeur x pour l'individu ω .

Si l'individu est choisi au hasard, la variable est dite aléatoire (v.a.).



Si, en Mathématiques, l'usage est plutôt d'appeler f, g ou h les applications, en Statistique celles-ci sont notées X, Y ou Z . Les minuscules x, y ou z représentent alors les réalisations (valeurs) de ces variables (applications). On fera donc bien la différence entre majuscules et minuscules pour éviter toute confusion entre applications et valeurs.

Dans toute la suite du cours, les v.a. considérées seront quantitatives.



Une variable aléatoire possède une loi de probabilité ^a qui régit son comportement.

Si la variable est discrète, la loi est définie par un diagramme en bâtons.

Si la variable est continue, le diagramme en bâtons est remplacé par une courbe de densité.

a. Si la Statistique fournit des informations sur une population à partir d'observations, les Probabilités fournissent des modèles théoriques pour étudier toute situation avec une part d'aléatoire.

Définition (n -échantillon aléatoire).

Un n -échantillon aléatoire est un n -uplet de v.a. (X_1, \dots, X_n) où X_i est le caractère X du i -ème individu choisi au hasard.

Définition (n -échantillon aléatoire simple).

Un n -échantillon aléatoire (X_1, \dots, X_n) est dit simple si les v.a. sont indépendantes.

Cela se produit si les individus sont choisis au hasard :

- soit avec remise
- soit sans remise (ou simultanément) à condition que le taux de sondage $\frac{n}{N}$ soit inférieur à 10%.



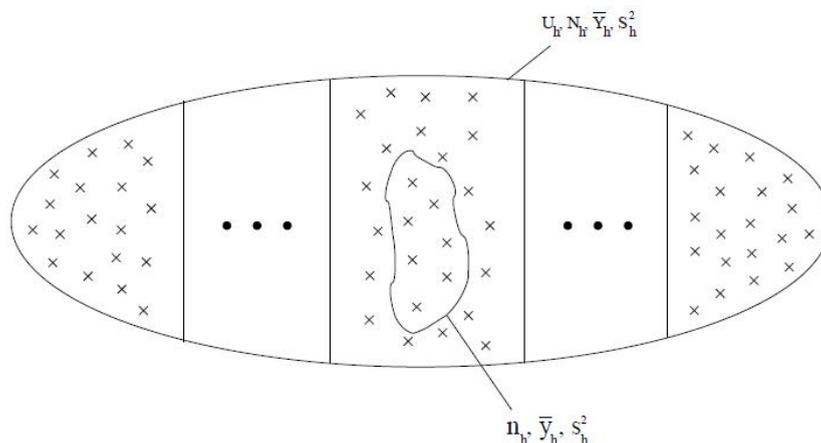
Le premier cas est théorique et le deuxième pratique.

2. Sondage en strates

Dans un sondage aléatoire simple, tous les échantillons d'une population de taille N sont possibles avec la même probabilité. On imagine que certains d'entre eux puissent s'avérer a priori indésirables.

Plus concrètement, dans l'étude du lancement d'un nouveau produit financier, on peut supposer des différences de comportement entre les "petits" et les "gros" clients de la banque. Il serait malencontreux que le hasard de l'échantillonnage conduise à n'interroger que les clients appartenant à une seule de ces catégories, ou simplement que l'échantillon soit trop déséquilibré en faveur de l'une d'elles. S'il existe dans la base de sondage une information auxiliaire permettant de distinguer, a priori, les catégories de "petits" et "gros" clients, on aura tout à gagner à utiliser cette information pour répartir l'échantillon dans chaque sous-population.

C'est le principe de la stratification : découper la population en sous-ensembles appelés strates et réaliser un sondage aléatoire simple dans chacune d'elles.



Nous avons alors deux manières de choisir les n_h :

- allocation proportionnelle : $\forall h, \frac{n_h}{N_h} = \frac{n}{N}$
- allocation optimale^[1] : $\forall h, \frac{n_h}{N_h} = n \frac{S_h}{\sum_h N_h S_h}$.

Dans toute la suite du cours, les échantillons aléatoires seront supposés simples.

III. Paramètres étudiés

1. Moyenne et variance d'une v.a.

On note X la v.a. étudiée.

On verra comment on peut inférer la loi de X , mais on s'intéressera avant tout aux informations suivantes :

- sa moyenne définie par :

$$m = E(X) = \begin{cases} \sum_i x_i P(X = x_i) & \text{si } X \text{ est discrète} \\ \int_{\mathbb{R}} x f(x) dx & \text{si } X \text{ est continue de densité } f \end{cases}$$

- sa variance définie par :

$$\sigma^2 = V(X) = E((X - m)^2) = \begin{cases} \sum_i (x_i - m)^2 P(X = x_i) & \text{si } X \text{ est discrète} \\ \int_{\mathbb{R}} (x - m)^2 f(x) dx & \text{si } X \text{ est continue de densité } f \end{cases}$$

[1]. On cherche la répartition de l'échantillon qui maximise la précision (et donc qui minimise la variance). Pour cela, on va augmenter les effectifs échantillonnés dans les strates où la variabilité est grande et diminuer les effectifs échantillonnés dans les strates homogènes.



- Rappelons les formules de Statistique descriptive pour les 3 types de données suivants :

1. Les données brutes (individu par individu) de la forme :

$$x_1, x_2, \dots, x_n$$

2. Les données regroupées par valeur, dans le cas discret, de la forme :

x_1	x_2	...	x_p
n_1	n_2	...	n_p

3. Les données regroupées par classe, dans le cas continu, de la forme :

$[e_1, e_2[$	$[e_2, e_3[$...	$[e_p, e_{p+1}[$
n_1	n_2	...	n_p

En notant $n = \sum_{i=1}^p n_i$ l'effectif total, $f_i = \frac{n_i}{n}$ la fréquence associée à n_i et $c_i = \frac{e_i + e_{i+1}}{2}$ le centre de la classe $[e_i, e_{i+1}[$, on a :

Type de données	1	2	3
Moyenne \bar{x}	$\frac{1}{n} \sum_{i=1}^n x_i$	$\frac{1}{n} \sum_{i=1}^p n_i x_i = \sum_{i=1}^p f_i x_i$	$\frac{1}{n} \sum_{i=1}^p n_i c_i = \sum_{i=1}^p f_i c_i$
Variance s^2	$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$	$\frac{1}{n} \sum_{i=1}^p n_i (x_i - \bar{x})^2 = \sum_{i=1}^p f_i (x_i - \bar{x})^2$	$\frac{1}{n} \sum_{i=1}^p n_i (c_i - \bar{x})^2 = \sum_{i=1}^p f_i (c_i - \bar{x})^2$

- Les formules théoriques dans le cas discret sont à comparer avec celles de la colonne 2!

2. Proportion (hors programme)

On s'intéressera à la proportion p c'est à dire à la part des individus dans une population possédant un certain caractère.



p est aussi la moyenne de la v.a. de Bernoulli qui, à un individu, associe 1 s'il possède le caractère désiré et 0 sinon. Les résultats sur p se déduiront donc de ceux sur m en prenant $X \sim \mathcal{B}(p)$.

Chapitre B

Estimation

Sommaire

I. Estimation ponctuelle	11
1. Estimateur et Loi Forte des Grands Nombres (LFGN)	11
2. Qualités d'un estimateur	12
II. Estimation par intervalle de confiance	14
1. Principe	14
2. Moyenne	14
3. Variance	16
4. Proportion (hors programme)	17

L'estimation consiste à donner une valeur approchée (ou un ensemble de valeurs plausibles) du paramètre **inconnu** m , σ^2 ou p , ceci à l'aide d'un échantillon de n observations issues de la population.

On considère alors :

- X la v.a. étudiée
- θ le paramètre inconnu à estimer : sa moyenne m ou sa variance σ^2
- (X_1, \dots, X_n) un échantillon

I. Estimation ponctuelle

1. Estimateur et Loi Forte des Grands Nombres (LFGN)

Définition (estimateur convergent).

Un estimateur convergent de θ est une fonction de X_1, \dots, X_n qui converge vers θ :

$$T_n = \phi(X_1, \dots, X_n) \xrightarrow[n \rightarrow +\infty]{} \theta$$

En pratique, on retiendra que $T_n \simeq \theta$ pour n assez grand (supérieur ou égal à 30).



- Pour être plus rigoureux (pas nécessaire pour nous), il faudrait préciser le type de convergence :
 - s'il s'agit d'une convergence en probabilité a , l'estimateur est dit convergent
 - s'il s'agit d'une convergence presque sûre b , l'estimateur est dit fortement convergent
- Pour ne pas alourdir les notations, on notera T au lieu de T_n

a. $T_n \xrightarrow[n \rightarrow +\infty]{\mathcal{P}} \theta$ si et seulement si $\forall \epsilon > 0, \mathcal{P}(|T_n - \theta| \geq \epsilon) \xrightarrow[n \rightarrow +\infty]{} 0$

b. $T_n \xrightarrow[n \rightarrow +\infty]{p.s.} \theta$ si et seulement si $\forall c \in C, T_n(c) \xrightarrow[n \rightarrow +\infty]{} \theta$ pour une certaine partie C de Ω vérifiant $\mathcal{P}(C) = 1$

Dans toute la suite du cours, un estimateur désignera un estimateur (fortement) convergent.

Théorème (LFGN).

La moyenne empirique converge (presque sûrement) vers la moyenne théorique :

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[n \rightarrow +\infty]{(p.s.)} m$$

En pratique, on retiendra que $\frac{1}{n} \sum_{i=1}^n X_i \simeq m$ pour n assez grand (supérieur ou égal à 30).

On en déduit :

Propriété (estimateurs usuels).

- La moyenne empirique $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ est un estimateur de m .
On le notera plus simplement \bar{X} .
- La variance empirique $S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \left(\frac{1}{n} \sum_{i=1}^n X_i^2 \right) - \left(\frac{1}{n} \sum_{i=1}^n X_i \right)^2$ est un estimateur de σ^2 .
On le notera plus simplement S^2 .
- La fréquence empirique $F_n = \frac{1}{n} \sum_{i=1}^n X_i$ est un estimateur de p où $X \sim \mathcal{B}(p)$.
On le notera plus simplement F .



Les formules empiriques sont à comparer, cette fois, avec celles de la colonne 1 du tableau page 8!

Définition (estimation ponctuelle).

Une estimation ponctuelle de θ est une réalisation d'un estimateur de θ .



Ne pas confondre l'estimateur \bar{X} (en majuscule) qui est une v.a. et l'estimation \bar{x} (en minuscule) qui est une valeur.

2. Qualités d'un estimateur

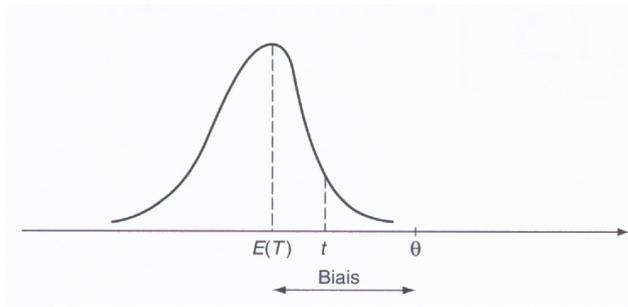
On considère ici T un estimateur de θ .

Définition (erreur - décomposition fluctuation d'échantillonnage / biais).

L'**erreur d'estimation** $T - \theta$ est une v.a. qui se décompose de la manière suivante :

$$T - \theta = (T - E(T)) + (E(T) - \theta)$$

- $T - E(T)$ est une erreur aléatoire : T varie autour de sa valeur centrale $E(T)$.
Cette partie de l'erreur est appelée **fluctuation d'échantillonnage**.
- $E(T) - \theta$ est une erreur systématique : T varie autour de sa valeur centrale $E(T)$ et non autour de θ .
Cette partie de l'erreur est appelée **biais**.



Il est souhaitable d'utiliser des estimateurs **sans biais** c'est à dire vérifiant $E(T) = \theta$.



1. Montrer que \bar{X} est un estimateur sans biais de m .
2. Montrer que S^2 est un estimateur biaisé^a qui a tendance à sous-estimer σ^2 .
3. Montrer que S^2 est un estimateur **asymptotiquement sans biais** de σ^2 c'est à dire vérifiant $E(S^2) \xrightarrow{n \rightarrow +\infty} \sigma^2$.

a. On pourra d'abord montrer que $S^2 = \left[\frac{1}{n} \sum_{i=1}^n (X_i - m)^2 \right] - (\bar{X} - m)^2$

Propriété (variance empirique corrigée).

La variance empirique corrigée $S^{*2} = \frac{n}{n-1} S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ est un estimateur sans biais de σ^2 .



L'écart-type empirique corrigé S^* reste biaisé pour σ mais asymptotiquement sans biais.

Propriété (erreur quadratique moyenne - décomposition biais / variance).

La précision de T est mesurée à l'aide de l'erreur quadratique moyenne $E((T - \theta)^2)$ qui se décompose sous la forme :

$$E((T - \theta)^2) = (E(T) - \theta)^2 + V(T)$$



L'objectif est de minimiser cette erreur quadratique moyenne.

Entre deux estimateurs sans biais du même paramètre, on choisira celui de plus petite variance^a.

a. L'Estimateur Sans Biais de Variance Minimale (ESBVM) sort du cadre de ce cours, tout comme l'Estimateur du Maximum de Vraisemblance (EMV)!



On considère $D = \frac{1}{n} \sum_{i=1}^n (X_i - m)^2$ et on admet qu'il s'agit d'un estimateur sans biais de σ^2 .

1. Montrer que $V(D) = \frac{1}{n}(\mu_4 - \sigma^4)$ où $\mu_4 = E((X - m)^4)$ désigne le moment centré d'ordre 4 de X .
2. Montrer que $V(S^{*2}) = \frac{1}{n}(\mu_4 - \frac{n-3}{n-1}\sigma^4)$.
3. Conclure.

II. Estimation par intervalle de confiance

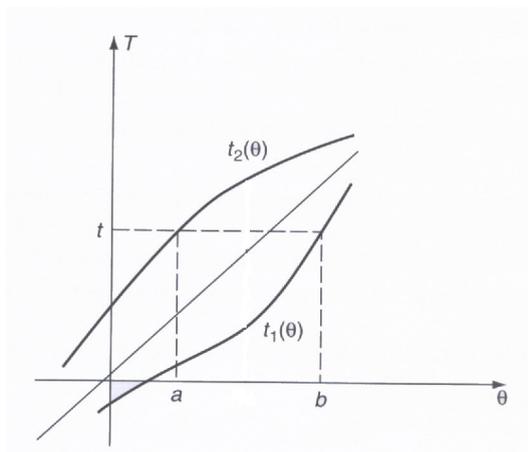
Plutôt que de fournir un renseignement du type $\theta \simeq c$, il est souvent plus intéressant de fournir un renseignement du type $a < \theta < b$ qui est certes moins précis, mais qui a l'avantage d'être accompagné d'une confiance.

1. Principe

Soit T un estimateur de θ (le meilleur possible) dont on connaît la loi de probabilité qui est fonction de θ .

Pour déterminer un intervalle de confiance pour θ au niveau $1 - \alpha$ c'est à dire un intervalle $[a, b]$ avec une chance de contenir θ égale à $1 - \alpha$ ou encore un risque de ne pas le contenir égal à α , il suffit de déterminer un intervalle de probabilité pour T au niveau $1 - \alpha$ c'est à dire deux réels $t_1(\theta) < t_2(\theta)$ vérifiant :

$$P\left(t_1(\theta) < T < t_2(\theta)\right) = 1 - \alpha$$



On lit alors l'intervalle de confiance $[a, b]$ selon l'horizontale issue de t (réalisation de T)



- Si on augmente le niveau de confiance $1 - \alpha$, les courbes s'écartent et donc l'intervalle grandit
- Si la taille de l'échantillon augmente, les courbes se rapprochent et donc l'intervalle diminue

2. Moyenne

Supposons $X \sim \mathcal{N}(m, \sigma)$ et estimons m .

Quand σ est connu

On utilise \bar{X} le meilleur estimateur de m , et la fonction pivotale ^[1] $W = \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1)$ (savez-vous démontrer ce résultat?).

L'intervalle de probabilité (à risques symétriques) pour W au niveau $1 - \alpha$ est :

$$-u_{(1-\frac{\alpha}{2})} < W < u_{(1-\frac{\alpha}{2})}$$

où $u_{(1-\frac{\alpha}{2})}$ désigne le quantile d'ordre $(1 - \frac{\alpha}{2})$ de la $\mathcal{N}(0, 1)$ ^[2].

L'intervalle de probabilité (à risques symétriques) pour \bar{X} au niveau $1 - \alpha$ est donc :

$$m - u_{(1-\frac{\alpha}{2})} \frac{\sigma}{\sqrt{n}} < \bar{X} < m + u_{(1-\frac{\alpha}{2})} \frac{\sigma}{\sqrt{n}}$$

L'intervalle de confiance ^[3] (bilatéral ^[4]) pour m au niveau $1 - \alpha$ est alors :

$$\bar{x} - u_{(1-\frac{\alpha}{2})} \frac{\sigma}{\sqrt{n}} < m < \bar{x} + u_{(1-\frac{\alpha}{2})} \frac{\sigma}{\sqrt{n}}$$

[1]. Il s'agit d'une fonction de X_1, \dots, X_n qui dépend de θ , mais dont la loi ne dépend pas de θ

[2]. C'est à dire le nombre pour lequel l'aire sous la courbe à gauche vaut $(1 - \frac{\alpha}{2})$. Par exemple, $u_{0.975} = 1,96$.

[3]. Certains auteurs fournissent plutôt l'intervalle de confiance "aléatoire" : $\bar{X} - u_{(1-\frac{\alpha}{2})} \frac{\sigma}{\sqrt{n}} < m < \bar{X} + u_{(1-\frac{\alpha}{2})} \frac{\sigma}{\sqrt{n}}$

[4]. On ne considérera pas l'intervalle de confiance unilatéral à droite (resp. gauche) qui regroupe du côté gauche (resp. droit) le risque α de ne pas contenir le paramètre et montrer ainsi que ce dernier est suffisamment grand (resp. petit)

Quand σ est inconnu

On utilise encore \bar{X} , mais cette fois la fonction pivotale $W = \frac{\bar{X} - m}{\frac{S^*}{\sqrt{n}}} \sim \mathcal{T}_{n-1}$ [5].

L'intervalle de probabilité pour W au niveau $1 - \alpha$ est :

$$-t_{(1-\frac{\alpha}{2})} < W < t_{(1-\frac{\alpha}{2})}$$

où $t_{(1-\frac{\alpha}{2})}$ désigne le quantile d'ordre $(1 - \frac{\alpha}{2})$ de la \mathcal{T}_{n-1} .

L'intervalle de probabilité pour \bar{X} au niveau $1 - \alpha$ est donc :

$$m - t_{(1-\frac{\alpha}{2})} \frac{S^*}{\sqrt{n}} < \bar{X} < m + t_{(1-\frac{\alpha}{2})} \frac{S^*}{\sqrt{n}}$$

L'intervalle de confiance pour m au niveau $1 - \alpha$ est alors :

$$\bar{x} - t_{(1-\frac{\alpha}{2})} \frac{s^*}{\sqrt{n}} < m < \bar{x} + t_{(1-\frac{\alpha}{2})} \frac{s^*}{\sqrt{n}}$$

Quand l'échantillon n'est plus gaussien mais de grande taille

On utilise le Théorème Central Limite :

Théorème (TCL).

Toujours en notant $m = E(X)$ et $\sigma^2 = V(X)$, on a :

$$\frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

En pratique, on retiendra que la loi de $\frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}}$ est proche de la gaussienne centrée réduite pour $n \geq 30$.

Quand σ est connu, le TCL nous assure :

$$W = \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

Autrement dit, la fonction asymptotiquement pivotale $W = \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}}$ suit approximativement une $\mathcal{N}(0, 1)$ et donc ...

Quand σ est inconnu, le TCL accompagné du théorème de Slutsky nous assure :

$$W = \frac{\bar{X} - m}{\frac{S^*}{\sqrt{n}}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

Autrement dit, la fonction asymptotiquement pivotale $W = \frac{\bar{X} - m}{\frac{S^*}{\sqrt{n}}}$ suit approximativement une $\mathcal{N}(0, 1)$ et donc ...

Les **intervalles de confiance** qui en découlent sont qualifiés d'**asymptotiques**.



Un artisan qui fabrique des objets de maroquinerie souhaite estimer le nombre moyen m de porte-cartes vendus quotidiennement. En notant ses ventes sur 36 jours, il obtient une moyenne de 120 et un écart-type corrigé de 17.

Donner un intervalle de confiance pour m au niveau 95% dans les deux cas suivants :

1. Si le nombre de porte-cartes est gaussien.
2. Sans l'hypothèse de normalité.

[5]. $\mathcal{T}_n = \frac{U}{\sqrt{\frac{X}{n}}}$ avec $U \sim \mathcal{N}(0, 1)$ et $X \sim \chi_n^2$ indépendantes (cf. plus loin pour la définition de χ_n^2)

3. Variance

Supposons $X \sim \mathcal{N}(m, \sigma)$ et estimons σ^2 .

Quand m est connue

On utilise D le meilleur estimateur de σ^2 , et la fonction pivotale $W = \frac{nD}{\sigma^2} \sim \chi_n^2$ [6].

L'intervalle de probabilité pour W au niveau $1 - \alpha$ est :

$$k_{\frac{\alpha}{2}} < \frac{nD}{\sigma^2} < k_{(1-\frac{\alpha}{2})}$$

où $k_{\frac{\alpha}{2}}, k_{(1-\frac{\alpha}{2})}$ désignent les quantiles d'ordre $\frac{\alpha}{2}$ et $(1 - \frac{\alpha}{2})$ de la χ_n^2 .

L'intervalle de probabilité pour D au niveau $1 - \alpha$ est donc :

$$k_{\frac{\alpha}{2}} \frac{\sigma^2}{n} < D < k_{(1-\frac{\alpha}{2})} \frac{\sigma^2}{n}$$

L'intervalle de confiance pour σ^2 au niveau $1 - \alpha$ est alors :

$$\boxed{\frac{nd}{k_{(1-\frac{\alpha}{2})}} < \sigma^2 < \frac{nd}{k_{\frac{\alpha}{2}}}}$$

Quand m est inconnue

On utilise cette fois S^2 comme estimateur de σ^2 , et la fonction pivotale $W = \frac{nS^2}{\sigma^2} \sim \chi_{n-1}^2$.

L'intervalle de probabilité pour W au niveau $1 - \alpha$ est :

$$k_{\frac{\alpha}{2}} < \frac{nS^2}{\sigma^2} < k_{(1-\frac{\alpha}{2})}$$

où $k_{\frac{\alpha}{2}}, k_{(1-\frac{\alpha}{2})}$ désignent cette fois [7] les quantiles d'ordre $\frac{\alpha}{2}$ et $(1 - \frac{\alpha}{2})$ de la χ_{n-1}^2 .

L'intervalle de probabilité pour S^2 au niveau $1 - \alpha$ est donc :

$$k_{\frac{\alpha}{2}} \frac{\sigma^2}{n} < S^2 < k_{(1-\frac{\alpha}{2})} \frac{\sigma^2}{n}$$

L'intervalle de confiance pour σ^2 au niveau $1 - \alpha$ est alors :

$$\boxed{\frac{ns^2}{k_{(1-\frac{\alpha}{2})}} < \sigma^2 < \frac{ns^2}{k_{\frac{\alpha}{2}}}}$$

Quand l'échantillon n'est plus gaussien mais de grande taille

Si la loi de X est unimodale pas trop dissymétrique, on utilise la fonction asympt. pivotale $W = \frac{S^{*2} - \sigma^2}{\sqrt{\frac{2S^{*4}}{n-1}}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$



En mesurant la quantité d'alcool (gr/l) contenue dans 10 cidres doux du marché, on obtient :

5,42 – 5,55 – 5,61 – 5,91 – 5,93 – 6,15 – 6,20 – 6,79 – 7,07 – 7,37

Estimer la variance par intervalle de confiance au niveau 95% en supposant la quantité d'alcool gaussienne.

[6]. $\chi_n^2 = \sum_{i=1}^n U_i^2$ avec les $U_i \sim \mathcal{N}(0, 1)$ indépendantes

[7]. On ne précise pas le degré de liberté pour ne pas alourdir les notations donc attention...

4. Proportion (hors programme)

Supposons l'échantillon de grande taille et estimons p .

Le nombre d'individus nF possédant le caractère étudié dans l'échantillon suit une loi binomiale $\mathcal{B}(n, p)$ donc, si n est grand, l'approximation d'une binomiale par une gaussienne^[8] fournit :

$$nF \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}\left(np, \sqrt{np(1-p)}\right)$$

ou encore :

$$F \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$$

On utilise alors F le meilleur estimateur de p , et la fonction asymptotiquement pivotale $W = \frac{F - p}{\sqrt{\frac{p(1-p)}{n}}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$.

L'intervalle de probabilité (asymptotique) pour F au niveau $1 - \alpha$ est donc :

$$p - u_{(1-\frac{\alpha}{2})} \sqrt{\frac{p(1-p)}{n}} < F < p + u_{(1-\frac{\alpha}{2})} \sqrt{\frac{p(1-p)}{n}}$$

L'intervalle de confiance (asymptotique) pour p au niveau $1 - \alpha$ est alors^[9] :

$$f - u_{(1-\frac{\alpha}{2})} \sqrt{\frac{f(1-f)}{n}} < p < f + u_{(1-\frac{\alpha}{2})} \sqrt{\frac{f(1-f)}{n}}$$



Il s'agit d'un intervalle de confiance asymptotique



Un échantillon de 100 votants choisis au hasard parmi tous les votants d'une circonscription a montré que 55% d'entre eux étaient favorables à un certain candidat.

1. Estimer la proportion de votants favorables à ce candidat par intervalle de confiance au niveau 95%.
2. Déterminer la taille de l'échantillon minimal pour assurer, au niveau 95%, une incertitude^a n'excédant pas 0,02.

a. Il s'agit de la demi-longueur de l'intervalle de confiance $f \pm u_{(1-\frac{\alpha}{2})} \sqrt{\frac{f(1-f)}{n}}$ c'est à dire $u_{(1-\frac{\alpha}{2})} \sqrt{\frac{f(1-f)}{n}}$

[8]. Le TCL s'exprime aussi sous la forme :

$$\bar{X} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}\left(m, \frac{\sigma}{\sqrt{n}}\right)$$

ou encore :

$$\sum_{i=1}^n X_i \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}\left(nm, \sqrt{n\sigma^2}\right)$$

Dans le cas où $X \sim \mathcal{B}(p)$, on obtient donc l'approximation d'une binomiale par une gaussienne :

$$\mathcal{B}(n, p) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}\left(np, \sqrt{np(1-p)}\right)$$

[9]. Admis (utilisation des coniques)

Chapitre C

Tests d'hypothèses

Sommaire

I. Introduction	19
1. Les faiseurs de pluie	19
2. Principe	21
II. Tests de conformité	21
1. Moyenne	21
2. Variance	22
3. Proportion (hors programme)	23
III. Tests de comparaison de deux échantillons indépendants (projet)	23
1. Moyennes	23
2. Variances	24
3. Proportions (hors programme)	26

I. Introduction

1. Les faiseurs de pluie

Des relevés effectués pendant de nombreuses années ont permis d'établir que le niveau naturel des pluies dans la Beauce en millimètres par an suit une loi normale $\mathcal{N}(600, 100)$. Des entrepreneurs, surnommés faiseurs de pluie, prétendaient pouvoir augmenter de 50 mm le niveau moyen de pluie, ceci par insémination des nuages au moyen d'iodure d'argent. Leur procédé fut mis à l'essai entre 1951 et 1959 et on releva les hauteurs de pluies suivantes :

Année	1951	1952	1953	1954	1955	1956	1957	1958	1959
mm	510	614	780	512	501	534	603	788	650

Deux hypothèses s'affrontaient :

- ou bien l'insémination était sans effet
- ou bien elle augmentait réellement le niveau moyen de pluie de 50 mm

Si m désigne la moyenne de la v.a. X égale au niveau annuel de pluie, ces **hypothèses** pouvaient se formaliser comme suit :

$$\begin{cases} H_0 : m = 600 \text{ mm} \\ H_1 : m = 650 \text{ mm} \end{cases}$$

Les agriculteurs hésitaient à opter pour le procédé onéreux des faiseurs de pluie, ils avaient besoin d'être convaincus...



| Comment pouvaient-ils décider?

Puisqu'il s'agit de tester la valeur de m , il est naturel d'utiliser la moyenne empirique $\bar{X} \sim \mathcal{N}\left(m, \frac{\sigma}{\sqrt{n}}\right)$.

Si H_0 est vraie, alors $\bar{X} \sim \mathcal{N}\left(m_0, \frac{\sigma}{\sqrt{n}}\right)$ donc $W = \frac{\bar{X} - m_0}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1)$ et donc $P(W \geq 1,64) = 5\%$.

La règle de décision^[1] est alors :

- Si $w \geq 1,64$, on rejette H_0 (au profit de H_1) avec un risque de se tromper égal à $\alpha = 5\%$
- Si $w < 1,64$, on conserve H_0 (faute de preuves suffisantes)

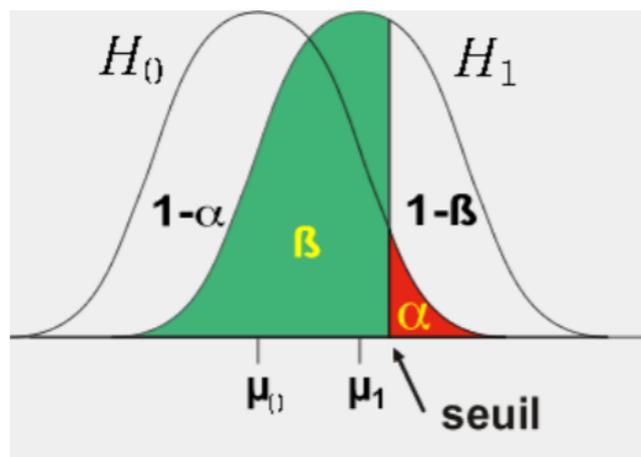
La v.a. W est appelée **variable de décision** et l'événement $\{W \geq 1,64\}$ la **région critique**.

Ici, les données relevées indiquent que $w = \frac{610,2 - 600}{100/3} = 0,306$ donc on conserve H_0 (aucune raison de rejeter H_0).



Le risque de se tromper en conservant H_0 est :

$$\beta = P\left(U < \frac{655 - 650}{100/3}\right) = P(U < 0,15) = 0,56 \quad \text{où } U \sim \mathcal{N}(0,1)$$



- Ces erreurs correspondent à des risques différents :
 - le risque de première espèce α consiste à acheter un procédé d'insémination inefficace
 - le risque de deuxième espèce β consiste à perdre une occasion d'augmenter le niveau de pluie
- Dans la pratique des tests statistiques, on se fixe α ce qui fait jouer à H_0 un rôle prééminent :
 - H_0 peut être une hypothèse solidement établie n'ayant jamais été contredite par l'expérience
 - H_0 peut être une hypothèse de prudence (l'innocuité d'un vaccin, l'innocence d'une personne)
 - H_0 peut être une hypothèse à laquelle on tient pour certaines raisons
- α étant fixé, β sera alors déterminé comme résultat d'un calcul (à condition que la loi de probabilité sous H_1 soit connue). Notons cependant que β varie dans le sens contraire de α . En effet, diminuer α conduit à une règle de décision plus stricte qui aboutit à rejeter H_0 que dans des cas rarissimes et donc à conserver H_0 bien souvent à tort ce qui revient à augmenter β ou encore à diminuer la puissance du test^a $1 - \beta$.

^a. La méthode de Neyman et Pearson permet de maximiser la puissance du test $1 - \beta$ pour une valeur donnée de α

[1]. Ce raisonnement probabiliste est à comparer avec le raisonnement par l'absurde sauf que le résultat impossible est ici remplacé par un résultat très peu probable, et la négation de l'hypothèse de départ par l'hypothèse alternative H_1

2. Principe

Un test est un mécanisme qui permet de trancher entre deux hypothèses à l'aide d'un échantillon.

En notant H_0 et H_1 ces deux hypothèses^[2], les quatre cas possibles sont représentés dans le tableau suivant :

Décision \ Vérité	H_0	H_1
	H_0	$1-\alpha$
H_1	α	$1-\beta$



On commence toujours par se fixer le risque de première espèce α .

Voici alors la démarche d'un test :

1. Choix de H_0 et H_1
2. Détermination de la variable de décision et de sa loi sous H_0
3. Détermination de la forme de la région critique (selon H_1) et de ses bornes (selon α)
4. Calcul de la valeur expérimentale de la variable de décision
5. Décision : rejet ou non de H_0

II. Tests de conformité

1. Moyenne

Supposons $X \sim \mathcal{N}(m, \sigma)$ et testons m .

Quand σ est connu

On utilise la variable de décision^[3] $W = \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}}$ qui, sous $H_0 : m = m_0$, vérifie :

$$W = \frac{\bar{X} - m_0}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1)$$



- Pour le test **unilatéral à droite** $\begin{cases} H_0 : m = m_0 \\ H_1 : m = m_1 \text{ avec } m_1 > m_0 \end{cases}$, la région critique est $[u_{(1-\alpha)}; +\infty[$.
- Pour le test unilatéral à droite $\begin{cases} H_0 : m = m_0 \\ H_1 : m > m_0 \end{cases}$, la région critique est encore $[u_{(1-\alpha)}; +\infty[$.
- $H_1 : m = m_1$ avec $m_1 > m_0$ est une hypothèse dite **simple** alors que $H_1 : m > m_0$ est une hypothèse dite **composite**.
- Pour le test **unilatéral à gauche** $\begin{cases} H_0 : m = m_0 \\ H_1 : m < m_0 \end{cases}$, la région critique est $] -\infty; -u_{(1-\alpha)}]$.
- Pour le test **bilatéral** $\begin{cases} H_0 : m = m_0 \\ H_1 : m \neq m_0 \end{cases}$, la région critique est $] -\infty; -u_{(1-\frac{\alpha}{2})}] \cup [u_{(1-\frac{\alpha}{2})}; +\infty[$.
- On fera essentiellement des tests unilatéraux avec H_1 composite.

[2]. On suppose qu'une et une seule est vraie

[3]. A comparer avec la fonction pivotale utilisée pour l'estimation par intervalle de confiance!

Quand σ est inconnu

On utilise la variable de décision $W = \frac{\bar{X} - m}{\frac{S^*}{\sqrt{n}}}$ qui, sous $H_0 : m = m_0$, vérifie :

$$W = \frac{\bar{X} - m_0}{\frac{S^*}{\sqrt{n}}} \sim \mathcal{T}_{n-1}$$

Quand l'échantillon n'est plus gaussien mais de grande taille

Quand σ est connu, on utilise la variable de décision $W = \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}}$ qui, sous $H_0 : m = m_0$, vérifie :

$$W = \frac{\bar{X} - m_0}{\frac{\sigma}{\sqrt{n}}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

Quand σ est inconnu, on utilise la variable de décision $W = \frac{\bar{X} - m}{\frac{S^*}{\sqrt{n}}}$ qui, sous $H_0 : m = m_0$, vérifie :

$$W = \frac{\bar{X} - m_0}{\frac{S^*}{\sqrt{n}}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$$



Lors d'une enquête sur le temps de sommeil par nuit des enfants de 2 à 3 ans dans un département français, on a trouvé une moyenne de 10,2 heures dans un groupe de 40 enfants avec un écart type de 2,1 heures. En France, la moyenne du temps de sommeil par nuit est de 11,7 heures chez les enfants de cet âge.

Est-il possible de conclure que le temps de sommeil moyen dans ce département est inférieur à celui de la France?

2. Variance

Supposons $X \sim \mathcal{N}(m, \sigma)$ et testons σ^2 .

Quand m est connue

On utilise la variable de décision $W = \frac{nD}{\sigma_0^2}$ qui, sous $H_0 : \sigma^2 = \sigma_0^2$, vérifie :

$$W = \frac{nD}{\sigma_0^2} \sim \chi_n^2$$

Quand m est inconnue

On utilise la variable de décision $W = \frac{nS^2}{\sigma_0^2}$ qui, sous $H_0 : \sigma^2 = \sigma_0^2$, vérifie :

$$W = \frac{nS^2}{\sigma_0^2} \sim \chi_{n-1}^2$$

Quand l'échantillon n'est plus gaussien mais de grande taille

Si la loi de X est unimodale pas trop dissymétrique, on utilise la var. de décision $W = \frac{S^{*2} - \sigma^2}{\sqrt{\frac{2S^{*4}}{n-1}}}$ qui, sous $H_0 : \sigma^2 = \sigma_0^2$, vérifie :

$$W = \frac{S^{*2} - \sigma_0^2}{\sqrt{\frac{2S^{*4}}{n-1}}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$$



Un agent immobilier prétend, lors d'une interview, que le prix moyen des transactions immobilières dans un quartier niçois est de 2400 euros du mètre carré avec un écart-type de 220 euros. Le journaliste chargé du dossier à paraître dans une revue spécialisée décide de vérifier ces affirmations à partir des 50 dernières transactions effectuées par quatre agences du quartier

1695	2202	2722	2534	2494	2648	2298	1997	2118	2767
2867	2391	2029	2121	2105	2565	2652	2497	2822	2713
2014	2350	2343	2398	2505	2630	2169	2661	2325	2031
2683	2328	2710	2417	2264	2299	2531	2423	2592	2577
2568	1992	2872	2603	2415	2072	2475	2089	2140	2720

On en tire : $\bar{x} = 2408.66$ et $s^* = 272.01$

En admettant que le prix de vente suive une loi normale, tester au seuil de 5% les affirmations de l'agent immobilier.

3. Proportion (hors programme)

Supposons l'échantillon de grande taille et testons p .

On utilise la variable de décision $W = \frac{F - p}{\sqrt{\frac{p(1-p)}{n}}}$ qui, sous $H_0 : p = p_0$, vérifie :

$$W = \frac{F - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$$



Sur un échantillon de 300 patients traités par un certain remède, 243 ont été guéris. La proportion de guérison est-elle significativement supérieure à 75% au seuil de 5%?

III. Tests de comparaison de deux échantillons indépendants (projet)

1. Moyennes

Supposons $X_A \sim \mathcal{N}(m_A, \sigma_A)$ et $X_B \sim \mathcal{N}(m_B, \sigma_B)$ indépendantes, et comparons m_A et m_B .

Quand les variances sont connues

On utilise la variable de décision $W = \frac{(\bar{X}_A - \bar{X}_B) - (m_A - m_B)}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}}$ qui, sous $H_0 : m_A = m_B$, vérifie :

$$W = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} \sim \mathcal{N}(0, 1)$$

Quand les variances sont inconnues et supposées égales

On utilise la variable de décision $W = \frac{(\bar{X}_A - \bar{X}_B) - (m_A - m_B)}{\sqrt{\frac{S_p^2}{n_A} + \frac{S_p^2}{n_B}}}$ qui, sous $H_0 : m_A = m_B$, vérifie :

$$W = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{S_p^2}{n_A} + \frac{S_p^2}{n_B}}} \sim \mathcal{T}_{n_A + n_B - 2}$$

où $S_p^2 = \frac{(n_A - 1)S_A^{*2} + (n_B - 1)S_B^{*2}}{n_A + n_B - 2}$ est la variance de pool^[4].

[4]. S_p^2 , appelée aussi variance combinée, n'est rien d'autre que la moyenne des variances corrigées des échantillons, pondérées par les tailles des échantillons diminuées de 1

Quand les variances sont inconnues et supposées différentes

On utilise la variable de décision $W = \frac{(\bar{X}_A - \bar{X}_B) - (m_A - m_B)}{\sqrt{\frac{S_A^{*2}}{n_A} + \frac{S_B^{*2}}{n_B}}}$ qui, sous $H_0 : m_A = m_B$, vérifie :

$$W = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{S_A^{*2}}{n_A} + \frac{S_B^{*2}}{n_B}}} \sim \mathcal{F}_m$$

où $m = \frac{1}{\frac{c^2}{n_A-1} + \frac{(1-c)^2}{n_B-1}}$ avec $c = \frac{\frac{S_A^2}{n_A-1}}{\frac{S_A^2}{n_A-1} + \frac{S_B^2}{n_B-1}}$.

Quand les échantillons ne sont plus gaussiens mais de grandes tailles

Quand les variances sont connues

On utilise la variable de décision $W = \frac{(\bar{X}_A - \bar{X}_B) - (m_A - m_B)}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}}$ qui, sous $H_0 : m_A = m_B$, vérifie :

$$W = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

Quand les variances sont inconnues et supposées égales

On utilise la variable de décision $W = \frac{(\bar{X}_A - \bar{X}_B) - (m_A - m_B)}{\sqrt{\frac{S_p^2}{n_A} + \frac{S_p^2}{n_B}}}$ qui, sous $H_0 : m_A = m_B$, vérifie :

$$W = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{S_p^2}{n_A} + \frac{S_p^2}{n_B}}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

Quand les variances sont inconnues et supposées différentes

On utilise la variable de décision $W = \frac{(\bar{X}_A - \bar{X}_B) - (m_A - m_B)}{\sqrt{\frac{S_A^{*2}}{n_A} + \frac{S_B^{*2}}{n_B}}}$ qui, sous $H_0 : m_A = m_B$, vérifie :

$$W = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{S_A^{*2}}{n_A} + \frac{S_B^{*2}}{n_B}}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$$



Le test de comparaison des moyennes doit être précédé par celui des variances dès que les variances sont inconnues

2. Variances

Supposons $X_A \sim \mathcal{N}(m_A, \sigma_A)$ et $X_B \sim \mathcal{N}(m_B, \sigma_B)$ indépendantes avec $S_A^* > S_B^*$, et comparons σ_A^2 et σ_B^2 .

Quand les moyennes sont connues

On utilise la variable de décision $W = \frac{\frac{D_A}{\sigma_A^2}}{\frac{D_B}{\sigma_B^2}}$ qui, sous $H_0 : \sigma_A^2 / \sigma_B^2 = 1$, vérifie :

$$W = \frac{D_A}{D_B} \sim \mathcal{F}(n_A, n_B)^{[5]}$$

[5]. $\mathcal{F}(n, p) = \frac{\chi_n^2/n}{\chi_p^2/p}$

Quand les moyennes sont inconnues

On utilise la variable de décision $W = \frac{\frac{S_A^{*2}}{\sigma_A^2}}{\frac{S_B^{*2}}{\sigma_B^2}}$ qui, sous $H_0 : \sigma_A^2 / \sigma_B^2 = 1$, vérifie :

$$W = \frac{S_A^{*2}}{S_B^{*2}} \sim \mathcal{F}(n_A - 1, n_B - 1)$$



Le test étant ici $\begin{cases} H_0 : \sigma_A^2 / \sigma_B^2 = 1 \\ H_1 : \sigma_A^2 / \sigma_B^2 > 1 \end{cases}$, la région critique est toujours $[f_{(1-\alpha)}; +\infty[$.



Les QI de 9 enfants d'un quartier d'une grande ville ont une moyenne de 107 avec un écart-type de 10. Les QI de 12 enfants d'un autre quartier ont une moyenne de 112 avec un écart-type de 9. On suppose que la variable aléatoire associée au QI suit une loi Normale.

Y a-t-il une différence significative au seuil de 5% entre les QI moyens des 2 quartiers?



Quand les échantillons ne sont plus gaussiens mais de grandes tailles

Si les lois sont unimodales pas trop dissymétriques, on utilise $W = \frac{(S_A^{*2} - S_B^{*2}) - (\sigma_A^2 - \sigma_B^2)}{\sqrt{\frac{2S_A^{*4}}{n_A - 1} + \frac{2S_B^{*4}}{n_B - 1}}}$ qui, sous $H_0 : \sigma_A^2 = \sigma_B^2$, vérifie :

$$W = \frac{S_A^{*2} - S_B^{*2}}{\sqrt{\frac{2S_A^{*4}}{n_A - 1} + \frac{2S_B^{*4}}{n_B - 1}}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$$



Des essais cliniques sont menés auprès de 137 patients atteints d'une maladie pulmonaire sans gravité afin de tester l'efficacité d'un traitement à la pulmotrycine.

Le protocole est le suivant :

- Des exercices respiratoires sont prescrits à 67 patients choisis au hasard ainsi qu'un placebo (groupe témoin).
- Les mêmes exercices respiratoires sont prescrits aux 70 autres patients ainsi que de la pulmotrycine (groupe traité).
- Au bout de trois mois, l'amélioration de la capacité pulmonaire de chaque patient est mesurée sur une échelle de 0 (pas d'amélioration) à 10 (récupération totale).

Voici les résultats obtenus :

Amélioration	Groupe témoin (A)	Groupe traité (B)
0	2	0
1	8	0
2	4	3
3	7	0
4	14	10
5	9	14
6	5	13
7	4	17
8	7	10
9	2	3
10	5	0

On en tire : $\bar{x}_A = 4.78$, $s_A^2 = 7.37$, $\bar{x}_B = 6$ et $s_B^2 = 2.66$

L'amélioration moyenne du groupe traité est supérieure à celle du groupe témoin (de combien?) mais cette différence doit-elle être attribuée aux bienfaits de la pulmotrycine ou aux fluctuations d'échantillonnage (le même protocole sur d'autres individus n'aurait sans doute pas donné les mêmes résultats). Autrement dit, la supériorité observée entre les deux moyennes empiriques est-elle significative au seuil de 2%?

3. Proportions (hors programme)

Supposons les deux échantillons de grandes tailles et indépendants, et comparons p_A et p_B .

On utilise la variable de décision $W = \frac{(F_A - F_B) - (p_A - p_B)}{\sqrt{\hat{p}(1 - \hat{p})(\frac{1}{n_A} + \frac{1}{n_B})}}$ qui, sous $H_0 : p_A = p_B$, vérifie :

$$W = \frac{F_A - F_B}{\sqrt{\hat{p}(1 - \hat{p})(\frac{1}{n_A} + \frac{1}{n_B})}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

où $\hat{p} = \frac{n_A f_A + n_B f_B}{n_A + n_B}$.



Lors des primaires d'une campagne présidentielle, des sympathisants d'un parti sont interrogés sur leur opinion à propos d'un candidat avant et après un débat télévisé. Avant le débat, 64% des 980 personnes interrogées déclarent avoir une opinion positive sur le candidat. Après le débat, cette proportion n'est plus que de 61% chez 1001 autres personnes interrogées.

Cette baisse est-elle significative au seuil de 5%?



Il existe des tests pour comparer des échantillons dépendants...

Annexes

Table 3

Loi Normale Centrée Réduite

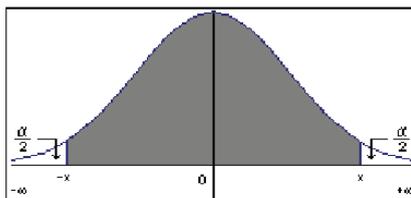
Fonction de répartition $F(z)=P(Z<z)$

Exemple : $P(Z<1.96)= 0.97500$ se trouve en ligne 1.9 et colonne 0.06

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,50000	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,52790	0,53188	0,53586
0,1	0,53983	0,54380	0,54776	0,55172	0,55567	0,55962	0,56356	0,56750	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59484	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,62930	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,65910	0,66276	0,66640	0,67003	0,67365	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69498	0,69847	0,70194	0,70540	0,70884	0,71226	0,71566	0,71904	0,72241
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,75490
0,7	0,75804	0,76115	0,76424	0,76731	0,77035	0,77337	0,77637	0,77935	0,78231	0,78524
0,8	0,78815	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0,9	0,81594	0,81859	0,82121	0,82382	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1,0	0,84135	0,84375	0,84614	0,84850	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,86650	0,86864	0,87076	0,87286	0,87493	0,87698	0,87900	0,88100	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90148
1,3	0,90320	0,90490	0,90658	0,90824	0,90988	0,91149	0,91309	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,92220	0,92364	0,92507	0,92647	0,92786	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93575	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,94520	0,94630	0,94738	0,94845	0,94950	0,95053	0,95154	0,95254	0,95352	0,95449
1,7	0,95544	0,95637	0,95728	0,95819	0,95907	0,95994	0,96080	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,97320	0,97381	0,97441	0,97500	0,97558	0,97615	0,97670
2,0	0,97725	0,97778	0,97831	0,97882	0,97933	0,97982	0,98030	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,98300	0,98341	0,98382	0,98422	0,98461	0,98500	0,98537	0,98574
2,2	0,98610	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,98840	0,98870	0,98899
2,3	0,98928	0,98956	0,98983	0,99010	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,99180	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,99430	0,99446	0,99461	0,99477	0,99492	0,99506	0,99520
2,6	0,99534	0,99547	0,99560	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,99720	0,99728	0,99736
2,8	0,99744	0,99752	0,99760	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861
3,0	0,99865	0,99869	0,99874	0,99878	0,99882	0,99886	0,99889	0,99893	0,99897	0,99900
3,1	0,99903	0,99906	0,99910	0,99913	0,99916	0,99918	0,99921	0,99924	0,99926	0,99929
3,2	0,99931	0,99934	0,99936	0,99938	0,99940	0,99942	0,99944	0,99946	0,99948	0,99950
3,3	0,99952	0,99953	0,99955	0,99957	0,99958	0,99960	0,99961	0,99962	0,99964	0,99965
3,4	0,99966	0,99968	0,99969	0,99970	0,99971	0,99972	0,99973	0,99974	0,99975	0,99976

Table 4

Loi de Student



α	1	0,8	0,6	0,4	0,2	0,1	0,05	0,02	0,01	0,002	0,001
$1 - \alpha$	0	0,2	0,4	0,6	0,8	0,9	0,95	0,98	0,99	0,998	0,999
$v = \text{ddl}$											
1	0,0000	0,3249	0,7265	1,3764	3,0777	6,3137	12,706	31,821	63,656	318,29	636,58
2	0,0000	0,2887	0,6172	1,0607	1,8856	2,9200	4,3027	6,9645	9,9250	22,328	31,600
3	0,0000	0,2767	0,5844	0,9785	1,6377	2,3534	3,1824	4,5407	5,8408	10,214	12,924
4	0,0000	0,2707	0,5686	0,9410	1,5332	2,1318	2,7765	3,7469	4,6041	7,1729	8,6101
5	0,0000	0,2672	0,5594	0,9195	1,4759	2,0150	2,5706	3,3649	4,0321	5,8935	6,8685
6	0,0000	0,2648	0,5534	0,9057	1,4398	1,9432	2,4469	3,1427	3,7074	5,2075	5,9587
7	0,0000	0,2632	0,5491	0,8960	1,4149	1,8946	2,3646	2,9979	3,4995	4,7853	5,4081
8	0,0000	0,2619	0,5459	0,8889	1,3968	1,8595	2,3060	2,8965	3,3554	4,5008	5,0414
9	0,0000	0,2610	0,5435	0,8834	1,3830	1,8331	2,2622	2,8214	3,2498	4,2969	4,7809
10	0,0000	0,2602	0,5415	0,8791	1,3722	1,8125	2,2281	2,7638	3,1693	4,1437	4,5868
11	0,0000	0,2596	0,5399	0,8755	1,3634	1,7959	2,2010	2,7181	3,1058	4,0248	4,4369
12	0,0000	0,2590	0,5386	0,8726	1,3562	1,7823	2,1788	2,6810	3,0545	3,9296	4,3178
13	0,0000	0,2586	0,5375	0,8702	1,3502	1,7709	2,1604	2,6503	3,0123	3,8520	4,2209
14	0,0000	0,2582	0,5366	0,8681	1,3450	1,7613	2,1448	2,6245	2,9768	3,7874	4,1403
15	0,0000	0,2579	0,5357	0,8662	1,3406	1,7531	2,1315	2,6025	2,9467	3,7329	4,0728
16	0,0000	0,2576	0,5350	0,8647	1,3368	1,7459	2,1199	2,5835	2,9208	3,6861	4,0149
17	0,0000	0,2573	0,5344	0,8633	1,3334	1,7396	2,1098	2,5669	2,8982	3,6458	3,9651
18	0,0000	0,2571	0,5338	0,8620	1,3304	1,7341	2,1009	2,5524	2,8784	3,6105	3,9217
19	0,0000	0,2569	0,5333	0,8610	1,3277	1,7291	2,0930	2,5395	2,8609	3,5793	3,8833
20	0,0000	0,2567	0,5329	0,8600	1,3253	1,7247	2,0860	2,5280	2,8453	3,5518	3,8496
21	0,0000	0,2566	0,5325	0,8591	1,3232	1,7207	2,0796	2,5176	2,8314	3,5271	3,8193
22	0,0000	0,2564	0,5321	0,8583	1,3212	1,7171	2,0739	2,5083	2,8188	3,5050	3,7922
23	0,0000	0,2563	0,5317	0,8575	1,3195	1,7139	2,0687	2,4999	2,8073	3,4850	3,7676
24	0,0000	0,2562	0,5314	0,8569	1,3178	1,7109	2,0639	2,4922	2,7970	3,4668	3,7454
25	0,0000	0,2561	0,5312	0,8562	1,3163	1,7081	2,0595	2,4851	2,7874	3,4502	3,7251
26	0,0000	0,2560	0,5309	0,8557	1,3150	1,7056	2,0555	2,4786	2,7787	3,4350	3,7067
27	0,0000	0,2559	0,5306	0,8551	1,3137	1,7033	2,0518	2,4727	2,7707	3,4210	3,6895
28	0,0000	0,2558	0,5304	0,8546	1,3125	1,7011	2,0484	2,4671	2,7633	3,4082	3,6739
29	0,0000	0,2557	0,5302	0,8542	1,3114	1,6991	2,0452	2,4620	2,7564	3,3963	3,6595
30	0,0000	0,2556	0,5300	0,8538	1,3104	1,6973	2,0423	2,4573	2,7500	3,3852	3,6460
40	0,0000	0,2550	0,5286	0,8507	1,3031	1,6839	2,0211	2,4233	2,7045	3,3069	3,5510
50	0,0000	0,2547	0,5278	0,8489	1,2987	1,6759	2,0086	2,4033	2,6778	3,2614	3,4960
60	0,0000	0,2545	0,5272	0,8477	1,2958	1,6706	2,0003	2,3901	2,6603	3,2317	3,4602
70	0,0000	0,2543	0,5268	0,8468	1,2938	1,6669	1,9944	2,3808	2,6479	3,2108	3,4350
80	0,0000	0,2542	0,5265	0,8461	1,2922	1,6641	1,9901	2,3739	2,6387	3,1952	3,4164
90	0,0000	0,2541	0,5263	0,8456	1,2910	1,6620	1,9867	2,3685	2,6316	3,1832	3,4019
100	0,0000	0,2540	0,5261	0,8452	1,2901	1,6602	1,9840	2,3642	2,6259	3,1738	3,3905
200	0,0000	0,2537	0,5252	0,8434	1,2858	1,6525	1,9719	2,3451	2,6006	3,1315	3,3398
∞	0,0000	0,2533	0,5244	0,8416	1,2816	1,6449	1,9600	2,3263	2,5758	3,0903	3,2906

Table 5

Loi du χ^2

$$P(\chi_v^2 \geq \chi_{v,\alpha}^2) = \alpha$$

$1 - \alpha$	0,001	0,005	0,01	0,025	0,05	0,1	0,5	0,9	0,95	0,975	0,99	0,995	0,999
α	0,999	0,995	0,99	0,975	0,95	0,9	0,5	0,1	0,05	0,025	0,01	0,005	0,001
$v = \text{ddl}$													
1	0,00	0,00	0,00	0,00	0,00	0,02	0,45	2,71	3,84	5,02	6,63	7,88	10,83
2	0,00	0,01	0,02	0,05	0,10	0,21	1,39	4,61	5,99	7,38	9,21	10,60	13,82
3	0,02	0,07	0,11	0,22	0,35	0,58	2,37	6,25	7,81	9,35	11,34	12,84	16,27
4	0,09	0,21	0,30	0,48	0,71	1,06	3,36	7,78	9,49	11,14	13,28	14,86	18,47
5	0,21	0,41	0,55	0,83	1,15	1,61	4,35	9,24	11,07	12,83	15,09	16,75	20,51
6	0,38	0,68	0,87	1,24	1,64	2,20	5,35	10,64	12,59	14,45	16,81	18,55	22,46
7	0,60	0,99	1,24	1,69	2,17	2,83	6,35	12,02	14,07	16,01	18,48	20,28	24,32
8	0,86	1,34	1,65	2,18	2,73	3,49	7,34	13,36	15,51	17,53	20,09	21,95	26,12
9	1,15	1,73	2,09	2,70	3,33	4,17	8,34	14,68	16,92	19,02	21,67	23,59	27,88
10	1,48	2,16	2,56	3,25	3,94	4,87	9,34	15,99	18,31	20,48	23,21	25,19	29,59
11	1,83	2,60	3,05	3,82	4,57	5,58	10,34	17,28	19,68	21,92	24,73	26,76	31,26
12	2,21	3,07	3,57	4,40	5,23	6,30	11,34	18,55	21,03	23,34	26,22	28,30	32,91
13	2,62	3,57	4,11	5,01	5,89	7,04	12,34	19,81	22,36	24,74	27,69	29,82	34,53
14	3,04	4,07	4,66	5,63	6,57	7,79	13,34	21,06	23,68	26,12	29,14	31,32	36,12
15	3,48	4,60	5,23	6,26	7,26	8,55	14,34	22,31	25,00	27,49	30,58	32,80	37,70
16	3,94	5,14	5,81	6,91	7,96	9,31	15,34	23,54	26,30	28,85	32,00	34,27	39,25
17	4,42	5,70	6,41	7,56	8,67	10,09	16,34	24,77	27,59	30,19	33,41	35,72	40,79
18	4,90	6,26	7,01	8,23	9,39	10,86	17,34	25,99	28,87	31,53	34,81	37,16	42,31
19	5,41	6,84	7,63	8,91	10,12	11,65	18,34	27,20	30,14	32,85	36,19	38,58	43,82
20	5,92	7,43	8,26	9,59	10,85	12,44	19,34	28,41	31,41	34,17	37,57	40,00	45,31
21	6,45	8,03	8,90	10,28	11,59	13,24	20,34	29,62	32,67	35,48	38,93	41,40	46,80
22	6,98	8,64	9,54	10,98	12,34	14,04	21,34	30,81	33,92	36,78	40,29	42,80	48,27
23	7,53	9,26	10,20	11,69	13,09	14,85	22,34	32,01	35,17	38,08	41,64	44,18	49,73
24	8,08	9,89	10,86	12,40	13,85	15,66	23,34	33,20	36,42	39,36	42,98	45,56	51,18
25	8,65	10,52	11,52	13,12	14,61	16,47	24,34	34,38	37,65	40,65	44,31	46,93	52,62
26	9,22	11,16	12,20	13,84	15,38	17,29	25,34	35,56	38,89	41,92	45,64	48,29	54,05
27	9,80	11,81	12,88	14,57	16,15	18,11	26,34	36,74	40,11	43,19	46,96	49,65	55,48
28	10,39	12,46	13,56	15,31	16,93	18,94	27,34	37,92	41,34	44,46	48,28	50,99	56,89
29	10,99	13,12	14,26	16,05	17,71	19,77	28,34	39,09	42,56	45,72	49,59	52,34	58,30
30	11,59	13,79	14,95	16,79	18,49	20,60	29,34	40,26	43,77	46,98	50,89	53,67	59,70

Pour $v > 30$, La loi du χ^2 peut être approximée par la loi normale $N(v, \sqrt{v})$

Table 6

Loi de Fisher F

$$P(F_{v_1, v_2} < f_{v_1, v_2, \alpha}) = \alpha$$

$\alpha = 0,975$

		v_1																	
		1	2	3	4	5	6	7	8	9	10	15	20	30	50	100	200	500	•
v_2	1	648	800	864	900	922	937	948	957	963	969	985	993	1001	1008	1013	1016	1017	1018
	2	38,5	39,0	39,2	39,3	39,3	39,4	39,4	39,4	39,4	39,4	39,4	39,4	39,4	39,5	39,5	39,5	39,5	39,5
	3	17,4	16,0	15,4	15,1	14,9	14,7	14,6	14,5	14,5	14,4	14,3	14,2	14,1	14,0	14,0	13,9	13,9	13,9
	4	12,2	10,6	9,98	9,60	9,36	9,20	9,07	8,98	8,90	8,84	8,66	8,56	8,46	8,38	8,32	8,29	8,27	8,26
	5	10,0	8,43	7,76	7,39	7,15	6,98	6,85	6,76	6,68	6,62	6,43	6,33	6,23	6,14	6,08	6,05	6,03	6,02
	6	8,81	7,26	6,60	6,23	5,99	5,82	5,70	5,60	5,52	5,46	5,27	5,17	5,07	4,98	4,92	4,88	4,86	4,85
	7	8,07	6,54	5,89	5,52	5,29	5,12	4,99	4,90	4,82	4,76	4,57	4,47	4,36	4,28	4,21	4,18	4,16	4,14
	8	7,57	6,06	5,42	5,05	4,82	4,65	4,53	4,43	4,36	4,30	4,10	4,00	3,89	3,81	3,74	3,70	3,68	3,67
	9	7,21	5,71	5,08	4,72	4,48	4,32	4,20	4,10	4,03	3,96	3,77	3,67	3,56	3,47	3,40	3,37	3,35	3,33
	10	6,94	5,46	4,83	4,47	4,24	4,07	3,95	3,85	3,78	3,72	3,52	3,42	3,31	3,22	3,15	3,12	3,09	3,08
	11	6,72	5,26	4,63	4,28	4,04	3,88	3,76	3,66	3,59	3,53	3,33	3,23	3,12	3,03	2,96	2,92	2,90	2,88
	12	6,55	5,10	4,47	4,12	3,89	3,73	3,61	3,51	3,44	3,37	3,18	3,07	2,96	2,87	2,80	2,76	2,74	2,72
	13	6,41	4,97	4,35	4,00	3,77	3,60	3,48	3,39	3,31	3,25	3,05	2,95	2,84	2,74	2,67	2,63	2,61	2,60
	14	6,30	4,86	4,24	3,89	3,66	3,50	3,38	3,29	3,21	3,15	2,95	2,84	2,73	2,64	2,56	2,53	2,50	2,49
	15	6,20	4,76	4,15	3,80	3,58	3,41	3,29	3,20	3,12	3,06	2,86	2,76	2,64	2,55	2,47	2,44	2,41	2,40
	16	6,12	4,69	4,08	3,73	3,50	3,34	3,22	3,12	3,05	2,99	2,79	2,68	2,57	2,47	2,40	2,36	2,33	2,32
	17	6,04	4,62	4,01	3,66	3,44	3,28	3,16	3,06	2,98	2,92	2,72	2,62	2,50	2,41	2,33	2,29	2,26	2,25
	18	5,98	4,56	3,95	3,61	3,38	3,22	3,10	3,01	2,93	2,87	2,67	2,56	2,44	2,35	2,27	2,23	2,20	2,19
	19	5,92	4,51	3,90	3,56	3,33	3,17	3,05	2,96	2,88	2,82	2,62	2,51	2,39	2,30	2,22	2,18	2,15	2,13
	20	5,87	4,46	3,86	3,51	3,29	3,13	3,01	2,91	2,84	2,77	2,57	2,46	2,35	2,25	2,17	2,13	2,10	2,09
22	5,79	4,38	3,78	3,44	3,22	3,05	2,93	2,84	2,76	2,70	2,50	2,39	2,27	2,17	2,09	2,05	2,02	2,00	
24	5,72	4,32	3,72	3,38	3,15	2,99	2,87	2,78	2,70	2,64	2,44	2,33	2,21	2,11	2,02	1,98	1,95	1,94	
26	5,66	4,27	3,67	3,33	3,10	2,94	2,82	2,73	2,65	2,59	2,39	2,28	2,16	2,05	1,97	1,92	1,90	1,88	
28	5,61	4,22	3,63	3,29	3,06	2,90	2,78	2,69	2,61	2,55	2,34	2,23	2,11	2,01	1,92	1,88	1,85	1,83	
30	5,57	4,18	3,59	3,25	3,03	2,87	2,75	2,65	2,57	2,51	2,31	2,20	2,07	1,97	1,88	1,84	1,81	1,79	
40	5,42	4,05	3,46	3,13	2,90	2,74	2,62	2,53	2,45	2,39	2,18	2,07	1,94	1,83	1,74	1,69	1,66	1,64	
50	5,34	3,98	3,39	3,06	2,83	2,67	2,55	2,46	2,38	2,32	2,11	1,99	1,87	1,75	1,66	1,60	1,57	1,55	
60	5,29	3,93	3,34	3,01	2,79	2,63	2,51	2,41	2,33	2,27	2,06	1,94	1,82	1,70	1,60	1,54	1,51	1,48	
80	5,22	3,86	3,28	2,95	2,73	2,57	2,45	2,36	2,28	2,21	2,00	1,88	1,75	1,63	1,53	1,47	1,43	1,40	
100	5,18	3,83	3,25	2,92	2,70	2,54	2,42	2,32	2,24	2,18	1,97	1,85	1,71	1,59	1,48	1,42	1,38	1,35	
200	5,10	3,76	3,18	2,85	2,63	2,47	2,35	2,26	2,18	2,11	1,90	1,78	1,64	1,51	1,39	1,32	1,27	1,23	
500	5,05	3,72	3,14	2,81	2,59	2,43	2,31	2,22	2,14	2,07	1,86	1,74	1,60	1,46	1,34	1,25	1,19	1,14	
•	5,02	3,69	3,12	2,79	2,57	2,41	2,29	2,19	2,11	2,05	1,83	1,71	1,57	1,43	1,30	1,21	1,13	1,00	

Loi de Fisher F (suite)

$$P(F_{v_1, v_2} < f_{v_1, v_2, \alpha}) = \alpha$$

$\alpha = 0,95$

		v_1																	
		1	2	3	4	5	6	7	8	9	10	15	20	30	50	100	200	500	•
v_2	1	161	200	216	225	230	234	237	239	241	242	246	248	250	252	253	254	254	254
	2	18,5	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5
	3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,70	8,66	8,62	8,58	8,55	8,54	8,53	8,53
	4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,86	5,80	5,75	5,70	5,66	5,65	5,64	5,63
	5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,62	4,56	4,50	4,44	4,41	4,39	4,37	4,37
	6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	3,94	3,87	3,81	3,75	3,71	3,69	3,68	3,67
	7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,51	3,44	3,38	3,32	3,27	3,25	3,24	3,23
	8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,22	3,15	3,08	3,02	2,97	2,95	2,94	2,93
	9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,01	2,94	2,86	2,80	2,76	2,73	2,72	2,71
	10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,85	2,77	2,70	2,64	2,59	2,56	2,55	2,54
	11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,72	2,65	2,57	2,51	2,46	2,43	2,42	2,40
	12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,62	2,54	2,47	2,40	2,35	2,32	2,31	2,30
	13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,53	2,46	2,38	2,31	2,26	2,23	2,22	2,21
	14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,46	2,39	2,31	2,24	2,19	2,16	2,14	2,13
	15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,40	2,33	2,25	2,18	2,12	2,10	2,08	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,35	2,28	2,19	2,12	2,07	2,04	2,02	2,01	
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,31	2,23	2,15	2,08	2,02	1,99	1,97	1,96	
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,27	2,19	2,11	2,04	1,98	1,95	1,93	1,92	
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,23	2,16	2,07	2,00	1,94	1,91	1,89	1,88	
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,20	2,12	2,04	1,97	1,91	1,88	1,86	1,84	
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,15	2,07	1,98	1,91	1,85	1,82	1,80	1,78	
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,11	2,03	1,94	1,86	1,80	1,77	1,75	1,73	
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,07	1,99	1,90	1,82	1,76	1,73	1,71	1,69	
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,04	1,96	1,87	1,79	1,73	1,69	1,67	1,65	
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,01	1,93	1,84	1,76	1,70	1,66	1,64	1,62	
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	1,92	1,84	1,74	1,66	1,59	1,55	1,53	1,51	
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,03	1,87	1,78	1,69	1,60	1,52	1,48	1,46	1,44	
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,84	1,75	1,65	1,56	1,48	1,44	1,41	1,39	
80	3,96	3,11	2,72	2,49	2,33	2,21	2,13	2,06	2,00	1,95	1,79	1,70	1,60	1,51	1,43	1,38	1,35	1,32	
100	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,97	1,93	1,77	1,68	1,57	1,48	1,39	1,34	1,31	1,28	
200	3,89	3,04	2,65	2,42	2,26	2,14	2,06	1,98	1,93	1,88	1,72	1,62	1,52	1,41	1,32	1,26	1,22	1,19	
500	3,86	3,01	2,62	2,39	2,23	2,12	2,03	1,96	1,90	1,85	1,69	1,59	1,48	1,38	1,28	1,21	1,16	1,11	
•	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,67	1,57	1,46	1,35	1,24	1,17	1,11	1,00	